A PARAMETRIC BOOTSTRAP USING THE FIRST FOURMOMENTS OF THE RESIDUALS

Pierre-Eric Treyens

GREQAM

Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III

Document de Travail
n°2007-11

A PARAMETRIC BOOTSTRAP USING THE FIRST FOUR MOMENTS OF THE RESIDUALS

Pierre-Eric TREYENS

Juin 2007

A parametric bootstrap using the first four moments of the residuals, and why the classical parametric bootstrap fails in a linear regression model.

Pierre-Eric TREYENS

Université de la Méditerranée
GREQAM

Université de Montréal
CIREQ

June 7, 2007

Abstract

We consider linear regression models and we suppose that disturbances are either Gaussian or nonGaussian. Until now, within the framework of the bootstrap, we thought that the error in rejection probability (ERP) had the same rate of convergence with the parametric bootstrap or the nonparametric bootstrap. For linear data generating processes (DGP) we show in this paper that this assertion is false if skewness and/or kurtosis coefficients of the distribution of the disturbances are nonnull. Indeed, we show that the ERP is the same for the asymptotic test as for the classical parametric bootstrap test it is based on. The only exception happens when we perform a t test or its associated bootstrap (parametric or not) in the model $y = \mu + \sigma u_t$ where the disturbances have nonnull kurtosis coefficient and a skewness coefficient equal to zero. In that case, the ERPs of any test we perform are of the same order. However, we provide a parametric bootstrap using the first four moments of the distribution of the residuals which is as accurate as a non parametric bootstrap which uses these first four moments implicitly. We will introduce it as the parametric bootstrap considering higher moments (CHM), and thus, we will speak about the parametric bootstrap CHM.

0 Introduction

Until now, it is conventional wisdom that bootstrap inferences, either parametric or nonparametric, are better than asymptotic ones, i.e. the error in the rejection probability (ERP) is smaller by using the bootstrap than by using asymptotic theory. It is clear that in order to decrease ERP, bootstrap must use extra information compared with the asymptotic test. So, the question is where this additional information comes from. We recall that a single bootstrap test may be based on a statistic τ in an asymptotic p-value form. Rejection by an asymptotic test at level α is then the event $\tau < \alpha$. Rejection by the bootstrap test is the event $\tau < Q(\alpha, \mu^*)$, where μ^* is the bootstrap data-generating process (DGP), and $Q(\alpha, \mu^*)$ is the (random) $\alpha-$quantile of the distribution of the statistic τ as generated by μ^*. Now, let us consider a bootstrap test computed from a t-statistic which tests $H_0 : \mu = \mu_0$ in a linear regression model where disturbances may not be Gaussian, which is the framework of this paper. By definition, we can write a t-statistic T as $T = \sqrt{n}\left(\frac{\hat{\mu}-\mu_0}{\hat{\sigma}_{\hat{\mu}}}\right)$ where n is the sample size and μ a parameter connected with any regressor of the linear regression model. So, computing a t-statistic just needs $\hat{\mu}$, the estimator of μ, and the estimator of its variance $\hat{\sigma}_{\hat{\mu}}$. Moreover, as the limit in distribution of a t-statistic is $N(0, 1)$, we can find an approximation of the CDF of T by using an Edgeworth expansion. And so, we can provide an approximation of the $\alpha-$quantile of T. Now, we just obtain the $\alpha-$quantile of the bootstrap distribution by replacing the true value of the higher moments of the disturbances by their random estimates as generated by the bootstrap DGP.

If we first consider a parametric bootstrap framework, the bootstrap DGP uses only the estimated variance of residuals $\hat{\sigma}$ which are directly connected to the estimated variance of the parameter tested. Indeed, bootstrap error terms are generated following a Gaussian distribution. So, it is clear we just use the same information as for computing the t-statistic. More precisely, the $\alpha-$quantile of the parametric bootstrap distribution just depends on the higher moments of a centered Gaussian random variable which are completely defined by its first two moments. So, the $\alpha-$quantile of the parametric bootstrap distribution is not random anymore. If we consider now a non parametric bootstrap, we implicitly use extra information which comes from higher moments of the distribution of the residuals. Indeed, when we resample the estimated residuals we provide random consistent estimators of these moments and so, the $\alpha-$quantile of the parametric bootstrap distribution is random. Obviously, if we introduce a new parametric bootstrap using estimated higher moments, we could provide a bootstrap framework which provides as much information as a non parametric bootstrap framework.

In the first part, we will introduce notations and definitions we will use later. In the following two parts, we will consider linear models $y = \mu_0 + \sigma u$ where we will test $\mu = \mu_0$, $y = \mu_0 + \beta_0 x + \sigma u$ where we will test $\beta = \beta_0$, and $y = \beta_0 x + \sigma u$ where we will test $\beta = \beta_0$. These three cases all include one-restriction tests which can occur in a linear regression model. For these three models, we are going to use a classical parametric bootstrap, a non parametric bootstrap and finally, a new parametric bootstrap which will use the estimators of the first four moments of the residuals.

This new parametric bootstrap will be called CHM parametric bootstrap, where CHM are the initials of "considering higher moments". In the last and fifth part, we will proceed to simulations and we will explain particularly how we apply this CHM parametric bootstrap.

1 Preliminaries

In this paper, we show how higher moments of the disturbances in a linear regression model influence either asymptotic and bootstrap inferences. In this way, we have to consider non Gaussian distributions whose skewness and/or kurtosis coefficients are not zero. For any centered random variable X, if we define its characteristic function f_c by $f_c(u) = E\left(e^{iux}\right)$ we can obtain, by using a MacLaurin development,

$$ln\left(f_c(u)\right) = \kappa_1(iu) + \frac{\kappa_2(iu)^2}{2!} + \frac{\kappa_3(iu)^3}{3!} + \cdots + \frac{\kappa_k(iu)^k}{k!} + \ldots \tag{1.1}$$

In this equation, the κ_k are order k cumulants of the distribution of X. Moreover, for a centered standardised random variable the four first cumulants are

$$\kappa_1 = 0,\ \kappa_2 = 1,\ \kappa_3 = E(X^3),\ \text{and}\ \kappa_4 = E(X^4) - 3 \tag{1.2}$$

In particular, κ_3 and κ_4 are the skewness and kurtosis coefficients. One of the main problems when we deal with higher moments is how we can generate centered standardised random variable fitting these coefficients. Treyens (2006) provides two methods to generate random variables in this way. Let us consider three independent random variables p, N_1 and N_2 where N_1 and N_2 are two Gaussian variables of expectations μ_1 and μ_2 and of standard error σ_1 and σ_2 and define $X = pN_1 + (1-p)N_2$. If p is a uniform distribution $U(0,1)$, the set of admissible couples (κ_3, κ_4) this method can provide is Γ as showed on the figure 1.1 and it will be called the unimodal method. If p is a binary variable and if $\frac{1}{2}$ is the probability that $p = 1$, the set of admissible couples is Γ_1 and this method will be called the bimodal method. On figure 1.1, the parabola and the straight line are just structural constraints which connect κ_4 to κ_3. Now, if a centered standardised random variable X has κ_3 and κ_4 as skewness and kurtosis coefficients, we will write $X \rightarrow \Delta(0, 1, \kappa_3, \kappa_4)$. In this paper, all disturbances will be distributed as $u_t \rightarrow ii\Delta(0, 1, \kappa_3, \kappa_4)$. To estimate the error in the rejection probability, we are going to use Edgeworth expansions. The main part of the asymptotic theory of bootstrap is based on Edgeworth expansions of statistics which follow asymptotically standard normal distributions. With this theory, we can express the error in the rejection probability as a quantity of the order of a negative power of n, where n is the size of the sample from which we compute the test statistic. Let t be a test statistic which asymptotically follows a standard normal distribution, and $F(.)$ be the CDF of the test statistic. Almost with a classic Taylor expansion, we can develop the function $F(.)$ as the CDF $\Phi(.)$ of the standard normal distribution plus an infinite sum of its successive derivatives that we always can write as a polynomial in t multiplied by the PDF $\phi(.)$ of the standard normal distribution. Precisely, we

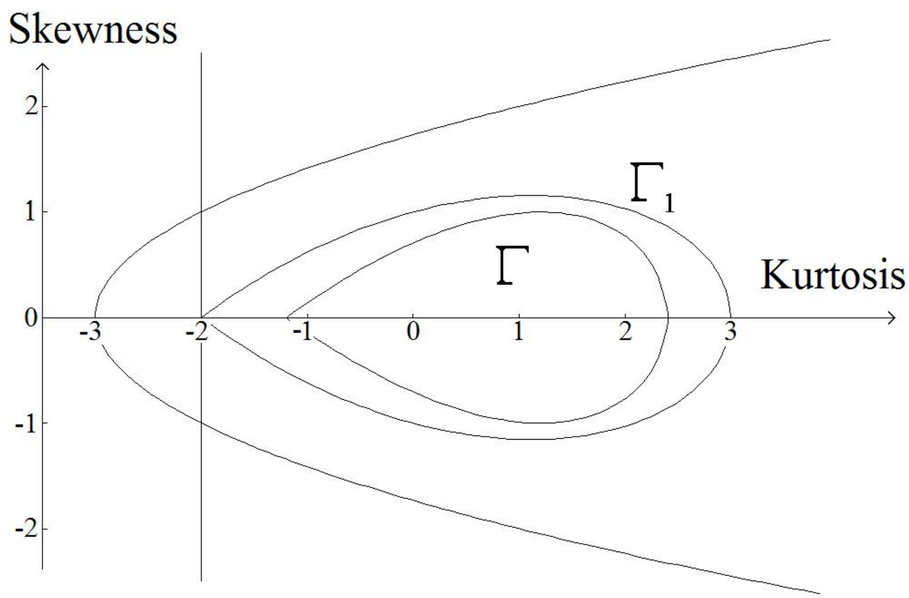

Figure 1.1: Sets of admissible couples Γ and Γ_1

have

$$F(t) = \Phi(t) - n^{-1/2}\phi(t)\sum_{i=1}^{\infty}\lambda_i He_{i-1}(t) \tag{1.3}$$

In this equation, $He_i(.)$ is the Hermite polynomial of degree i and the λ_i are coefficients which are at most of the order of unity. Hermite polynomials are implicitly defined by the relation $\phi^{(i)}(x) = (-1)^i He_i(x)\phi(x)$, as a function of the derivatives of $\phi(.)$ which gives the recurrence relation $He_0(x) = 1$ and $He_{i+1}(x) = xHe_i(x) - He_i^{'}(x)$, and the coefficients λ_i are defined as the following function of uncentered moments of the test statistic t, $\lambda_i = \frac{n^{1/2}}{i!}E\left(He_i(t)\right)$. Moreover, we will use in computations several random (or not) variables which are functions of the disturbances and of parameters of the models, all these variables (w, q, s, k, X, Q, m_1, m_3 and m_4) are described in the Appendix.

2 Testing a simple mean

Let us consider the model $y_t = \mu_0 + \sigma u_t$ with $u_t \to ii\Delta(0, 1, \kappa_3, \kappa_4)$. In order to test $H_0 : \mu = \mu_0$ we use a Student statistic and we bootstrap it. The t-statistic is obviously $T = \sqrt{n}\left(\frac{\hat{\mu}-\mu_0}{\hat{\sigma}_\mu}\right)$, where $\hat{\mu}$ is the estimator of the mean of the sample and $\hat{\sigma}$ is the unbiased estimator of the standard error of the OLS regression. So, we can give an approximate value at order $o_p\left(n^{-1}\right)$ of the test statistic T :

$$T = w\left(1 - n^{-1/2}\frac{q}{2} + n^{-1}\left(\frac{w^2}{2} - \frac{1}{2} + \frac{3q^2}{8}\right)\right) + o_p\left(n^{-1}\right) \tag{2.1}$$

In order to give the Edgeworth expansion $F_{1,T}(.)$ of the CDF of this test statistic, we have to check two points. First, the asymptotic distribution of the test statistic must be a standard normal distribution. Secondly, all expectations of its successive power must exist. The first point is easy to check, indeed, by applying the central limit theorem on w, we see that the asymptotic distribution of w is a standard normal distribution. Moreover, the limit in probability of the right hand of the equation 2.1 divided by w is deterministic and equal to 1. In order to check the second, we will just compute successive expectations of powers of T. This will allow us to deduce easily expectations of Hermite polynomials of T and in that way we obtain an estimate of $F_{1,T}(.)$ at order n^{-1}. Now, we can compute an approximation q_α of the $\alpha-$quantile $Q(\alpha, \Theta)$ of the test statistic at order n^{-1}, where Θ is the DGP which generated the original data y. To find this approximation, we introduce $q_\alpha = z_\alpha + n^{-1/2} q_{1\alpha} + n^{-1} q_{2\alpha}$, the Cornish-Fisher development of $Q(\alpha, \Theta)$, where z_α is the α-quantile of a standard normal distribution, with actually, $Q(\alpha, \Theta) = z_\alpha + n^{-1/2} q_{1\alpha} + n^{-1} q_{2\alpha} + o(n^{-1})$. If we now evaluate $F_{1,T}(.)$ in q_α, then we can find the expression of $q_{1\alpha}$ and $q_{2\alpha}$.

$$q_{1\alpha} = -\frac{\kappa_3 \left(1 + 2z_\alpha^2\right)}{6} \tag{2.2}$$

$$q_{2\alpha} = -\frac{z_\alpha^3 \left(6\kappa_4 - 20\kappa_3^2 - 18\right) + z_\alpha \left(-18\kappa_4 + 5\kappa_3^2 - 18\right)}{72} \tag{2.3}$$

And so, the ERP of the asymptotic test is obviously at order $n^{-1/2}$. Indeed, estimating $F_{1,T}(.)$ in z_α, we can provide the ERP of the asymptotic test

$$\begin{aligned} ERP_{as}^1 &= n^{-1/2}\phi(z_\alpha)\left[\frac{\kappa_3 \left(1 + 2z_\alpha^2\right)}{6}\right] \qquad (2.4) \\ &\quad + n^{-1}\phi(z_\alpha)\left[\frac{\kappa_3^2(3z_\alpha - 2z_\alpha^3 - z_\alpha^5)}{18} + \frac{\kappa_4(z_\alpha^3 - 3z_\alpha)}{12} - \frac{z_\alpha(1 + z_\alpha^2)}{4}\right] \end{aligned}$$

Now, to compute the ERP of a bootstrap test, we have first to find q_α^* the $\alpha-$quantile of the bootstrap statistic's distribution. So, we replace κ_3 by κ_3^* which is its estimate as generated by the bootstrap DGP, and we do not deal with κ_4 because it does not appear in $q_{1\alpha}$. The rejection condition of the bootstrap test is $T < q_\alpha^*$ at the order we consider and this condition is equivalent to $T - q_\alpha^* + q_\alpha < q_\alpha$. Now, we just have to compute the Edgeworth expansion $F^*(.)$ of $T - q_\alpha^* + q_\alpha$ to provide an estimate of its CDF and to eveluate it at q_α to find the ERP of the bootstrap test. If we consider a non parametric bootstrap or a parametric bootstrap CHM, we obtain exactly the same estimator $\hat{\kappa}_3$ of κ_3. Indeed, the first uses the empirical distribution of the residuals and the second the estimate of κ_3 provided by these residuals. Both these methods lead us to $\kappa_3^* = \kappa_3 + n^{-1/2}\left(s - 3w - \frac{3}{2}\kappa_3 q\right) + o_p\left(n^{-1/2}\right)$ which is random because s, w and q are. Then, we compute the Edgeworth expansion as described earlier and we obtain an ERP equal to zero at order $n^{-1/2}$ and nonnull at order n^{-1}. More precisely, we obtain

$$ERP_{BT_{nonpar}}^1 = n^{-1} z_\alpha \phi(z_\alpha) \frac{(1 + 2z_\alpha^2)(3\kappa_3^2 - 2\kappa_4)}{12} \tag{2.5}$$

Actually, Hall (1992) already obtained this result with a method quite different. Then, the parametric bootstrap DGP just uses a centered normal distribution to generate bootstrap samples. So, its third cumulant is zero and thus, it is not random. By using exactly the same method as previously, we obtain an ERP at order $n^{-1/2}$ as for an asymptotic test. Actually, this ERP is

$$\begin{aligned} ERP^1_{BT_{par}} &= -n^{-1/2}\phi(z_\alpha)\left[\frac{\kappa_3\left(1+2z_\alpha^2\right)}{6}\right] \\ &\quad +n^{-1}\phi(z_\alpha)\left[\frac{\kappa_4}{4}-\frac{7\kappa_3^2}{36}-\frac{\kappa_4 z_\alpha^2}{12}-\frac{\kappa_3^2 z_\alpha^4}{18}\right] \end{aligned} \tag{2.6}$$

So, for any κ_3 and κ_4, the ERP of the non parametric or of the CHM bootstrap test is at order n^{-1}. If the disturbances are Gaussian, $\kappa_3 = \kappa_4 = 0$ then the ERP of the non parametric bootstrap test or, in an equivalent way, of the bootstrap CHM is now at order $n^{-3/2}$. On the other hand, by considering 2.4 and 2.6, we see that if $\kappa_3 \neq 0$ then the dominant term of both ERP is the same and it is at order $n^{-1/2}$. Thus, if disturbances are asymmetrical, the parametric bootstrap fails to decrease ERP. However, if both κ_3 and κ_4 are null, i.e. if the first four moments of their distribution are the same as for a standard normal distribution, whichever bootstrap we use, then we obtain the same accuracy at order $n^{-3/2}$. Now if $\kappa_3 = 0$ and $\kappa_4 \neq 0$, the three tests have the same accuracy. This result is quite surprising, it just occurs because the ERP of the non parametric bootstrap test at order n^{-1} depends on the kurtosis coefficient κ_4. Indeed, intuitively, we thought that a non parametric bootstrap test was always better than the asymptotic test which it is based on. Another special case appears in equation 2.5 , when we have $3\kappa_3^2 = 2\kappa_4$, the ERP of the non parametric bootstrap test is now at order $n^{-1/2}$. In the next part we find this condition again and so, we will test it in the simulation part. In the next part, we will use a Student not on the intercept but on other variables. We will consider two cases, with or without an intercept in addition to this variable.

3 A linear model

3.1 With intercept

Now, we consider linear models $y_t^{,} = \mu_0^{,} + \alpha_0 x_t^{,} + Z_t\gamma + \sigma u_t^{,}$ with $u_t^{,} \rightarrow ii\Delta(0,1,\kappa_3,\kappa_4)$ and where Z is a $n \times k$ matrix. By projecting both the left and the right hand side of the defining equation of the model on $M_{\iota Z}$ and by using FWL theorem, we obtain the model $M_{\iota Z}y_t^{,} = \alpha M_{\iota Z}x_t^{,} + residuals$ with $\sum_{t=1}^n M_{\iota Z}y_t = \sum_{t=1}^n M_{\iota Z}x_t = 0$. Obviously, in this last model, if we want to test the null $H_0 : \alpha = \alpha_0$ the test statistic is a Student with $k+2$ degrees of freedom. In this part, we just consider the model $y_t = \mu_0 + \alpha x_t + \sigma u_t$ with $u_t \rightarrow ii\Delta(0,1,\kappa_3,\kappa_4)$. Or in an equivalent way, the model $y_t = \alpha x_t + \sigma\left(u_t - n^{-1/2}w\right)$ with $\sum_{t=1}^n y_t = \sum_{t=1}^n x_t = 0$ and two degrees of liberty. Moreover, we suppose that $Var(x) = 1$ without loss of generality. We obtain the asymptotic test

$$T = X\left(1 - n^{-1/2}\frac{q}{2} + n^{-1}\left(\frac{X^2}{2} + \frac{w^2}{2} - 1 + \frac{3q^2}{8}\right)\right) \tag{3.1}$$

The limit in probability of T is a standard normal distribution and we use Edgewoth expansions to provide an approximation of the CDF $F(.)$ of T at order n^{-1}. Following the same framework as in the previous part, we compute an approximation $q_\alpha = z_\alpha + n^{-1/2}q_{\alpha 1} + n^{-1}q_{\alpha 2}$ of the $\alpha-$quantile of T.

$$q_{\alpha 1} = \frac{\kappa_3 m_3\left(z_\alpha^2 - 1\right)}{6} \tag{3.2}$$

$$\begin{aligned} q_{\alpha 2} &= z_\alpha^3\left(\frac{(3\kappa_4+9)m_4 - 4\kappa_3^2 m_3^2 - 9\kappa_4 + 18}{72}\right) \\ &\quad + z_\alpha\left(\frac{(-9\kappa_4 - 27)m_4 + 10\kappa_3^2 m_3^2 + 27\kappa_4 + 18}{72}\right) \end{aligned} \tag{3.3}$$

And so, we obtain the ERP of the asymptotic test

$$ERP_{as}^2 = n^{-1/2}\kappa_3 m_3 \phi(z_\alpha)\left[\frac{1+z_\alpha^2}{6}\right] + o\left(n^{-\frac{1}{2}}\right) \tag{3.4}$$

We recall that the rejection condition at order n^{-1} of the bootstrap test is $T < q_\alpha^*$ where q_α^* is the approximation of the $\alpha-$quantile of the bootstrap distribution and now, we can write this condition as $T < q_\alpha^* - q_\alpha + q_\alpha$. In order to obtain q_α^*, we just replace κ_3 by its estimate as generated by the bootstrap DGP in $q_{\alpha 1}$ and $q_{\alpha 2}$.
If we deal with the non parametric or parametric bootstrap CHM, we obtain the same estimator as in the last part. We recall that $\kappa_3^* = \kappa_3 + n^{-1/2}\left(s - 3w - \frac{3}{2}\kappa_3 q\right) + o_p\left(n^{-1/2}\right)$. Now,we just use exactly the same framework as in the previous part and in this case, the CDF $F^*(.)$ of $T - q_\alpha^* + q_\alpha$ is the same as $F(.)$ the CDF of T at order n^{-1}. So when we evaluate $F^*(.)$ in q_α we obviously find an ERP at order $n^{-3/2}$. Intuitively, we thought we would find an ERP of the bootstrap test at order n^{-1} as in the previous part. But according to Davidson and MacKinnon (2000), independence between the test statistic and the bootstrap DGP improves bootstrap inferences by an $n^{-1/2}$ factor, this is the reason why we have $F(.) = F^*(.)$ up to order n^{-1} and why we find an ERP of the bootstrap test at order $n^{-3/2}$. Actually, we do not have independence but a weaker condition. Let B be the bootstrap DGP, it comes from the random part of κ_3^*, and so the random part of B is the same as the random part of κ_3^*. Here, we just have $E\left(T^k B\right) = o\left(n^{-1/2}\right)$ for all $k \in \aleph$. But this is enough for $F^*(.)$ to be equal to $F(.)$ at order n^{-1}. In fact, we obtain this result just because by introducing the intercept in linear model, we have $m_1 = 0$. Then, for the parametric bootstrap the estimator of κ_3 is still zero. We proceed in the same way as for the non parametric bootstrap or parametric bootstrap CHM and we obtain an ERP at order $n^{-1/2}$ which is exactly the same than ERP_{as}^2 as defined in equation 3.4. This result is natural, at least when we consider the order of the ERP, indeed we use as much information to perform asymptotic and parametric bootstrap tests, the estimators of

the parameter α and of the variance; and so, no more information, no more accuracy. Now, let us consider $\kappa_3 = 0$ and $\kappa_4 \neq 0$, i.e. symmetrical distributions for the disturbances. The ERP of both the non parametric bootstrap and parametric bootstrap CHM are still at order $n^{-3/2}$. This is logical, indeed whether $\kappa_3 = 0$ or not, the κ_3^* we use in the bootstrap DGP is a random variable with the true κ_3 as expectation. So, we do not use more information coming from the true DGP which generates the original data. However, the ERP of asymptotic and parametric bootstrap tests are now at order n^{-1} but they are no longer equal. Indeed, when $\kappa_3 = 0$ the ERP of the parametric bootstrap test is

$$ERP^2_{BT_{par}} = n^{-1}\phi\left(z_\alpha\right)\kappa_4\frac{\left(z_\alpha^3 - z_\alpha\right)\left(3 - m_4\right)}{24} + o\left(n^{-1}\right) \tag{3.5}$$

So the distribution of these two statistics are not the same, which we could have thought by only considering the case $\kappa_3 \neq 0$ and so, the parametric bootstrap still fails to improve the accuracy of inferences. The last case we have to consider is $\kappa_3 = \kappa_4 = 0$. Here, the parametric bootstrap estimates κ_3 and κ_4 perfectly because for a centered Gaussian distribution they are both equal to zero and it decreases its ERP at order $n^{-3/2}$ exactly as the non parametric bootstrap and parametric bootstrap CHM. Such a result just occurs because we use extra information by chance, using a bootstrap DGP very close to the original DGP. So, the parametric bootstrap test is better than the asymptotic test only when disturbances have skewness and kurtosis coefficients equal to zero whereas the non parametric bootstrap and parametric bootstrap CHM always improve the quality of the asymptotic t test. In particular, when disturbances are Gaussian, the parametric bootstrap has the same accuracy as the non parametric bootstrap and the bootstrap CHM.

3.2 Without intercept

Let us consider linear models $y_t^{\text{'}} = \lambda_0 x_t^{\text{'}} + Z_t\beta + \sigma^{\text{'}} v_t$ with $v_t \rightarrow ii\Delta(0, 1, \kappa_3^{\text{'}}, \kappa_4^{\text{'}})$ and Z a matrix $n \times k$. In order to test $H_0 : \lambda = \lambda_0$, we can use the FWL theorem to test this hypothesis in the model $M_Z y_t = \lambda M_Z x_t + residuals$ in an equivalent way. So, any Student test can be seen as a particular case of the Student test connected to α_0 in the model $y_t = \alpha_0 x_t + \sigma u_t$ with $u_t \rightarrow ii\Delta(0, 1, \kappa_3, \kappa_4)$ and with $k+1$ degrees of freedom rather than only one. Now, we just consider this last model with only one degree of freedomand where we impose $n^{-1}\sum_{t=1}^{n} x_t^2 = 1$ without loss of generality. The t statistic we compute is given by

$$T = X\left(1 - n^{-1/2}\frac{q}{2} + n^{-1}\left(\frac{X^2}{2} - \frac{1}{2} + \frac{3q^2}{8}\right)\right) \tag{3.6}$$

As the limit in probability of T is still a standard normal distribution, we can follow exactly the same procedure as in the previous part in order to obtain the approximation of the CDF $F(.)$ of T at order n^{-1} by using Edgeworth expansions and then an approximation $q_\alpha = z_\alpha + n^{-1/2}q_{\alpha 1} + n^{-1}q_{\alpha 2}$ at order n^{-1} of the $\alpha-$quantile of T.

Computations provide

$$q_{\alpha 1} = \frac{\left(\kappa_3 m_3 - 3\kappa_3 m_1\right) z_\alpha^2 - \kappa_3 m_3}{6} \tag{3.7}$$

$$\begin{aligned} q_{\alpha 2} &= z_\alpha^3 \left(\frac{\left(3\kappa_4 + 9\right) m_4 - 4\kappa_3^2 m_3^2 + 6\kappa_3^2 m_1 m_3 + 18\kappa_3^2 m_1^2 - 9\kappa_4 + 18}{72} \right) \\ &\quad + z_\alpha \left(\frac{\left(-9\kappa_4 - 27\right) m_4 + 10\kappa_3^2 m_3^2 - 6\kappa_3^2 m_1 m_3 - 9\kappa_3^2 m_1^2 + 27\kappa_4 - 18}{72} \right) \end{aligned} \tag{3.8}$$

As previously, the ERP of the asymptotic test T is at order $n^{-1/2}$ because $q_{\alpha 1} \neq 0$. Moreover, estimating $F(.)$ in z_α and if $\phi(.)$ is the PDF of a standard normal distribution, then we find that

$$ERP_{as}^3 = n^{-1/2} \kappa_3 \phi(z_\alpha) \left[\frac{m_3 + z_\alpha^2 (m_3 - 3m_1)}{6} \right] + o\left(n^{-\frac{1}{2}}\right) \tag{3.9}$$

Considering this last equation, we see that if $m_1 = 0$, we have $ERP_{as}^3 = ERP_{as}^2$ Now, whatever the bootstrap we consider, we have to center the residuals to provide a valid bootstrap DGP because the intercept does not belong to the model. We recall again that the rejection condition at order n^{-1} of the bootstrap test is $T < q_\alpha^*$ where q_α^* is the approximation of the $\alpha-$quantile of the bootstrap distribution and now, we can write this condition as $T < q_\alpha^* - q_\alpha + q_\alpha$. In order to obtain q_α^*, we just replace κ_3 by its estimate as generated by the bootstrap DGP in $q_{\alpha 1}$ and $q_{\alpha 2}$.
If we deal with the non parametric or parametric bootstrap CHM, we obtain the same estimator as in the last part. We recall that $\kappa_3^* = \kappa_3 + n^{-1/2}\left(s - 3w - \frac{3}{2}\kappa_3 q\right) + o_p\left(n^{-1/2}\right)$. Now,we just use exactly the same framework as in the previous part. In this part, we do not have anymore $m_1 = 0$, so we do not obtain $E\left(T^k B\right) = o\left(n^{-1/2}\right)$ for all $k \in \aleph$ and in this case, the CDF $F^*(.)$ of $T - q^* + q$ is not equal to $F(.)$ the CDF of T. Now, by estimating $F^*(.)$ in q_α, we find the ERP of both non parametric and parametric CHM bootstrap

$$ERP_{BT_{nonpar}}^3 = n^{-1} \phi\left(z_\alpha\right) \frac{m_1 z_\alpha \left(2\kappa_4 - 3\kappa_3^2\right) \left(m_3 \left(z_\alpha^2 - 1\right) - 3m_1 z_\alpha^2\right)}{12} + o\left(n^{-1}\right) \tag{3.10}$$

Considering last result, we see this ERP is at order $n^{-3/2}$ if $\kappa_4 = 0$. However, there is an other way to obtain this order for the ERP, this is the special case we already obtained by testing a simple mean in the previous chapter, we will deal with it in the simulation part in order to know if we can find this result again or if it is just a theoritical result. Now, let consider a parametric bootstrap test, the estimator of κ_3 used by bootstrap DGP is still zero, we proceed exactly in the same way as previously and we obtain an ERP which is the same than ERP_{as}^3 as defined in equation 3.9but when, distribution of the disturbances is symmetrical, i.e. when $\kappa_3 = 0$, we now have

$$ERP_{BT_{par}}^3 = n^{-1} \kappa_4 z_\alpha \frac{\left(m_4 - 3\right)\left(z_\alpha^2 - 3\right)}{24} + o\left(n^{-1}\right) \tag{3.11}$$

And now, explanations are the same as at the end of part 3.1

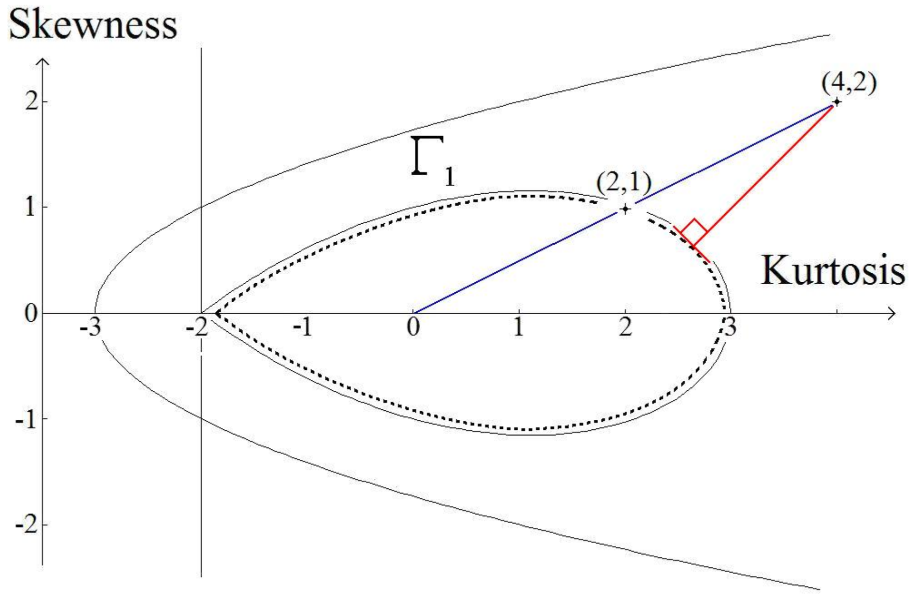

Figure 4.1: Methods of projection inside of Γ_1

4 Simulation evidence

In the different figures provided in appendix, we seek to estimate the power of these four tests and the level of significance is $\alpha = 0.05$. For the asymptotic test, there are 100000 repetitions and for the bootstrap tests we limit the number of repetitions to 20000 and bootstrap repetitions to 999. We want to examine the convergence rate when skewness and/or kurtosis coefficients of the distribution of the disturbances are varying in the set Γ_1. So we fit the kurtosis or the skewness with a specific value and we allow the other one to vary in Γ_1. Asymptotic tests, and parametric and non parametric bootstrap tests are performed in the usual way. However, in order to estimate the power of bootstrap CHM tests, a new problem arises in generating bootstrap samples. Indeed, even if we generate disturbances following a standard distribution belonging to the set Γ_1, then the estimated standardised residuals do not always provide an estimate $(\hat{\kappa}_4, \hat{\kappa}_3)$ which belongs to Γ_1. So, we cannot directly use the bimodal method to generate bootstrap samples. This problem happens because estimates of higher moments are not very reliable for small size samples. We correct it by multiplying $(\hat{\kappa}_4, \hat{\kappa}_3)$ by a constant $k \in [0, 1]$. In our algorithm, we choose $k = \frac{10-i}{10}$ with i the first integer in $[0, 10]$ which satisfies $(k\hat{\kappa}_4, k\hat{\kappa}_3) \in \Gamma_1 \setminus Fr(\Gamma_1)$. Actually, this homothetic transformation respects the signs of both estimated cumulants $\hat{\kappa}_3$ and $\hat{\kappa}_4$ and never provides a couple on the frontier of Γ_1. Indeed, on this frontier, the distributions connected with the couple (κ_4, κ_3) which defines it are not continuous. We provide an example on the figure 4.1 with $(\hat{\kappa}_4, \hat{\kappa}_3) = (4, 2)$, here we have $k = 5$ and we obtain the couple $(2, 1)$. In fact, we prefer this method rather than a method projecting directly on a subset very close to the frontier of Γ_1, as described on figure 4.1 because it projects in the direction of the cumulants of a standard normal distribution, i.e. $(\kappa_4, \kappa_3) = (0, 0)$. A last problem can occur when κ_3 is very close to 0, it is not a theoretical problem because solutions always exist in the set Γ_1 and it is just a computational problem. So, if $\hat{\kappa}_3 < 2.10^{-2}$, then we fit κ_3 to zero, in order to cancel all the algorithmic problems which can occur.

4.1 $y_t = \mu_0 + u_t$

Here, we suppose that $\mu_0 = 0$ and $u_t \sim iid(0, 1)$. Then, we test $H_0 : \mu = 0$ against $H_1 : \mu < 0$ because we deal with unilateral tests. We consider the following couples κ_3 and κ_4.

Couples $(\kappa_3; \kappa_4)$	$(0,8;0)$	$(0,4;0)$	$(0;0)$	$(-0,4;0)$	$(-0,8;0)$
Couples $(\kappa_3; \kappa_4)$	$(0,8;1)$	$(0,4;1)$	$(0;1)$	$(-0,4;1)$	$(-0,8;1)$

By considering figures 7.1 to 7.4, we check that parametric bootstrap tests and asymptotic ones provide the same rejection probabilities, in agreement with the theoretical results. In fact, even the sign of the ERP are the ones predicted by our computations. Then, as soon as $\kappa_3 \neq 0$, we check that asymptotic and parametric bootstrap tests have the same accuracy. Thus, we check that parametric bootstrap test does not use more information than asymptotic test when $\kappa_3 \neq 0$. Now, if we consider the next four figures, we first observe that non parametric bootstrap and parametric bootstrap CHM have the same convergence rates and they are better than parametric bootstrap or asymptotic tests. Thus, at the order we consider, we do not use more information than the one contained in the first four moments. Moreover, we can observe sub-reject and over-reject phenomenons, these are in agreement with theory. Actually, when $\kappa_3 > 0$, the tails of distributions are thicker on the left, so we have more chance to find a realization in the rejection area and to obtain an over-rejection probability. Then, when $\kappa_3 < 0$, it is exactly the reverse.

4.2 $y_t = \mu_0 + \alpha_0 x_t + u_t$

Here, we suppose that $\mu_0 = 2$ and $\alpha_0 = 0$ with $u_t \sim iid(0, 1)$ and $Var(x) = 1$. Then, we still have an unilateral test and we test $H_0 : \alpha = 0$ against $H_1 : \alpha < 0$. We consider the following couples κ_3 and κ_4.

Couples $(\kappa_3; \kappa_4)$	$(0,8;0)$	$(0,8;1)$	$(0;0)$	$(0;1)$

In this subsection, we observe quite the same results. When $\kappa_3 \neq 0$, convergence rates are less fast for both asymptotic and parametric bootstrap tests and they are the same in the both cases. On the other hand, non parametric bootstrap and parametric bootstrap CHM provide exactly the same results and theses two methods provide better convergence rates almost when κ_3 is very different from zero.

4.3 $y_t = \alpha_0 x_t + u_t$

Here, we suppose that $\alpha_0 = 0$ with $u_t \sim iid(0,1)$ and $Var(x) = 1$. Then, we still have an unilateral test and we test $H_0 : \alpha = 0$ against $H_1 : \alpha < 0$. We consider the following couples κ_3 and κ_4.

Couples $(\kappa_3;\kappa_4)$	$(0,8;0)$	$(0,8;1)$	$(0;0)$	$(0;1)$

Finally, in this last subsection, we still obtain the same results with convergence rates faster when $\kappa_3 = 0$ for asymptotic and parametric bootstrap tests than when $\kappa_3 = 1$. Moreover, convergence rates are the same for both methods. Then, for non parametric bootstrap and parametric bootstrap CHM, convergence rates are the same and we still observe over-reject when $\kappa_3 > 0$.

5 Conclusion

In this paper we introduced a new parametric bootstrap method which uses the four first moments of the estimated residuals. Asymptotically, this method has the same convergence rates as the non parametric bootstrap and they are better than asymptotic and parametric bootstrap since $\kappa_3 \neq 0$. Actually, the accuracy of a test is directly linked to the information it uses. As both asymptotic and parametric tests use the same information coming from the first two moments (except when $\kappa_3 = \kappa_4 = 0$), they provide the same convergence rates. On the other hand, non parametric bootstrap and parametric bootstrap CHM use extra-information from third and fourth moments and they provide better convergence rates. We resume the different results of this paper in the two following tabular.

Test	$\kappa_3 \neq 0 \neq \kappa_4$	$\kappa_3 = 0 \neq \kappa_4$	$\kappa_3 = \kappa_4 = 0$
Asymptotic	$O\left(n^{-\frac{1}{2}}\right)$	$O\left(n^{-1}\right)$	$O\left(n^{-1}\right)$
Parametric bootstrap	$O\left(n^{-\frac{1}{2}}\right)$	$O\left(n^{-1}\right)$	$O\left(n^{-\frac{3}{2}}\right)$
Non parametric bootstrap and CHM bootstrap	$O\left(n^{-1}\right)$	$O\left(n^{-1}\right)$	$O\left(n^{-\frac{3}{2}}\right)$

Tableau A. "Models $y_t = \mu_0 + \sigma_0 u_t$ and $y_t = \alpha_0 x_t + \sigma_0 u_t$"

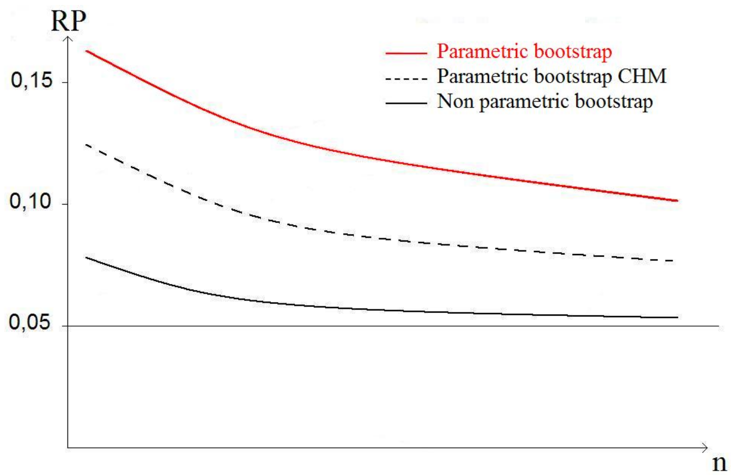

Figure 5.1: Rejection probabilities supposed

Test	$\kappa_3 \neq 0 \neq \kappa_4$	$\kappa_3 = 0 \neq \kappa_4$	$\kappa_3 = \kappa_4 = 0$
Asymptotic	$O\left(n^{-\frac{1}{2}}\right)$	$O\left(n^{-1}\right)$	$O\left(n^{-1}\right)$
Parametric bootstrap	$O\left(n^{-\frac{1}{2}}\right)$	$O\left(n^{-1}\right)$	$O\left(n^{-\frac{3}{2}}\right)$
Non parametric bootstrap and CHM bootstrap	$O\left(n^{-\frac{3}{2}}\right)$	$O\left(n^{-\frac{3}{2}}\right)$	$O\left(n^{-\frac{3}{2}}\right)$

Tableau B. "Model $y_t = \mu_0 + \alpha_0 x_t + \sigma_0 u_t$"

Then, even if we did not do the simulations, it seems logical to think that the results would be the same ones for any test in a linear regression model. Obviously, it could be very different for other models. Now, let us imagine an other model with rejection probabilities such as those in the figure 5.1. In this example, it would be obvious that other cumulants appear in the dominant term of the rejection probability that the first four cumulants. Actually, if we could develop other methods to control more than the first four cumulants of a distribution, it would be possible to know the information a bootstrap test uses because non parametric bootstrap always uses all the estimated moments of the residuals. Now, the obvious question is : "Does a non parametric bootstrap test always use the information contained in the first

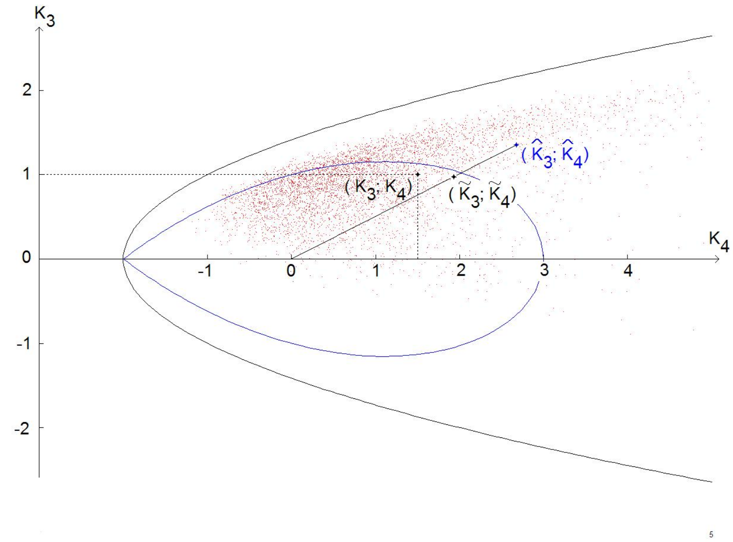

Figure 5.2: Distribution of couples $(\kappa_3; \kappa_4)$ estimated.

four moments ?" Parametric bootstrap CHM could help to answer this question. Then, our simulations show that if disturbances are normal, parametric bootstrap can provide better results than non parametric bootstrap or parametric bootstrap CHM, however, it is quite impossible for small samples to know if its distribution is normal or not. Moreover, even if parametric bootstrap CHM and non parametric bootstrap have almost the same rejection probabilities, the first can reject H_0 when the second does not and conversely. So, we think that we must use a principle of precaution using bootstrap and to compute the three bootstrap tests. To achieve this paper, we just show with the assistance of the figure 5.2 why the parametric bootstrap CHM can be more accurate than the non parametric bootstrap. On this figure, there are 5000 points which are estimated couples $(\hat{\kappa}_3, \hat{\kappa}_4)$ from a distribution whose $(\kappa_3, \kappa_4) = (1, 1.5)$. By considering this figure, we immediatly see that a lot of couples $(\hat{\kappa}_3, \hat{\kappa}_4)$ are apart from the set Γ_1 as defined on the figure 4.1. In these cases, parametric bootstrap CHM, by projecting towards the normality uses a PGD more close to the true distribution than the non parametric bootstrap. And so, parametric bootstrap CHM will provides better inferences than the non parametric bootstrap.

6 Bibliography

Davidson, R. and J. G. MacKinnon (1993). "Estimation and inference in econometrics", *New-York, Oxford University Press.*

Davidson, R. and J. G. MacKinnon (1999*b*). "The size distortion of bootstrap tests", *Econometric Theory*, **15**, 361-76.

Davidson, R. and J. G. MacKinnon (2000). "Improving the reliability of bootstrap test", *Working paper of Queen's University*, No. 995.

Davidson, R. and J. G. MacKinnon (2001). "The power of bootstrap and asymptotic tests", *Resampling Methods in Econometrics Conference*, Montreal University.

Davidson, R. and J. G. MacKinnon (2004). "Econometrics theory and methods", *New York, Oxford University Press.*

Hall P. (1992). "The bootstrap and the Edgeworth expansion", *New York, Springer Velag.*

Hill G. W. et Davis, A. W. (1968). "Generalized asymptotic expansions of Cornish-Fisher type", *Ann. Math. Statist. 39: 1264-1273.*

Treyens, P.-E. (2005). "Two methods to generate centered distributions controlling skewness and kurtosis coefficients", *Non soumis, version provisoire.*

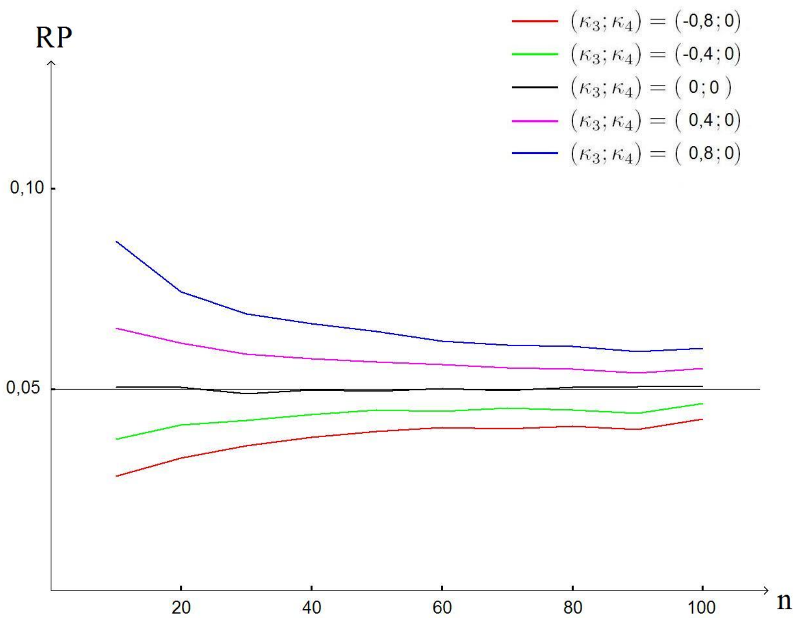

Figure 7.1: RP of asymptotic tests with $\kappa_4 = 0$ and κ_3 varying.

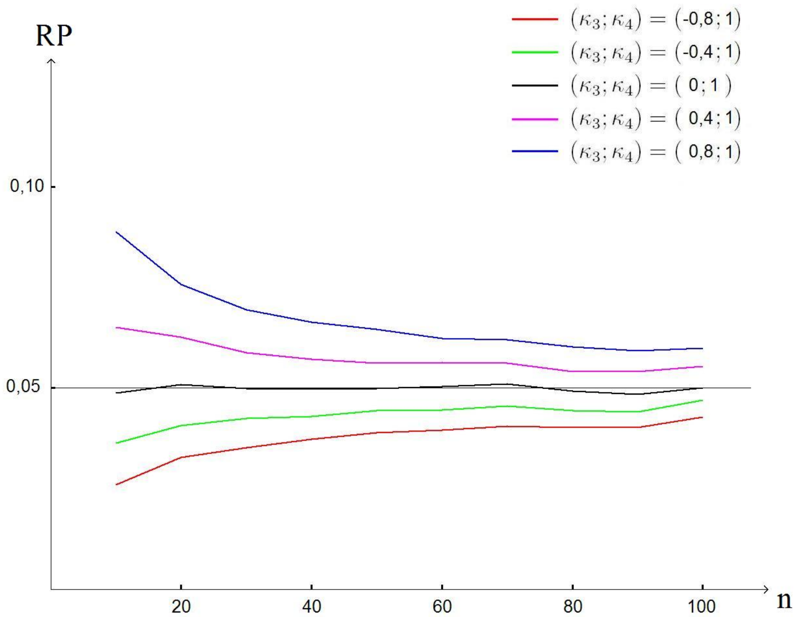

Figure 7.2: RP of asymptotic tests with $\kappa_4 = 1$ and κ_3 varying.

7 Appendix

7.1 $y_t = \mu_0 + u_t$

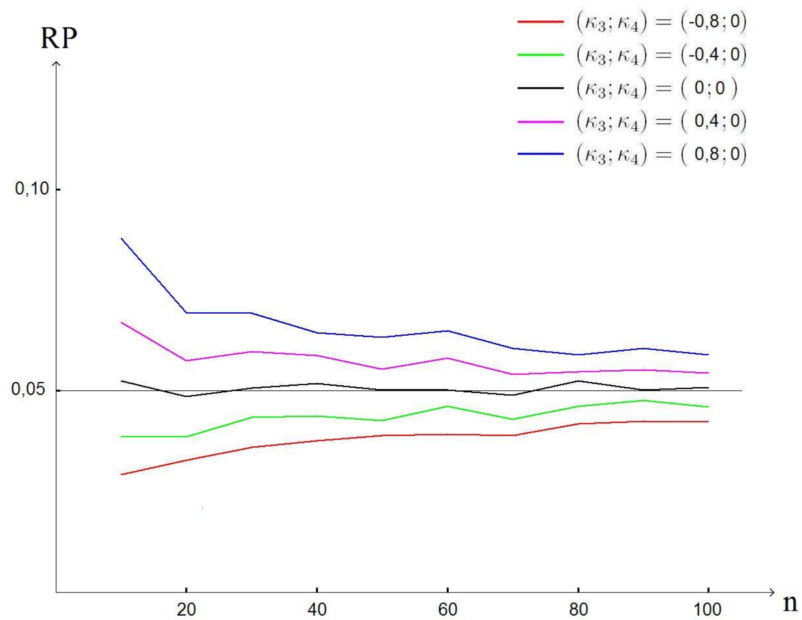

Figure 7.3: RP of parametric bootstrap tests with $\kappa_4 = 0$ and κ_3 varying.

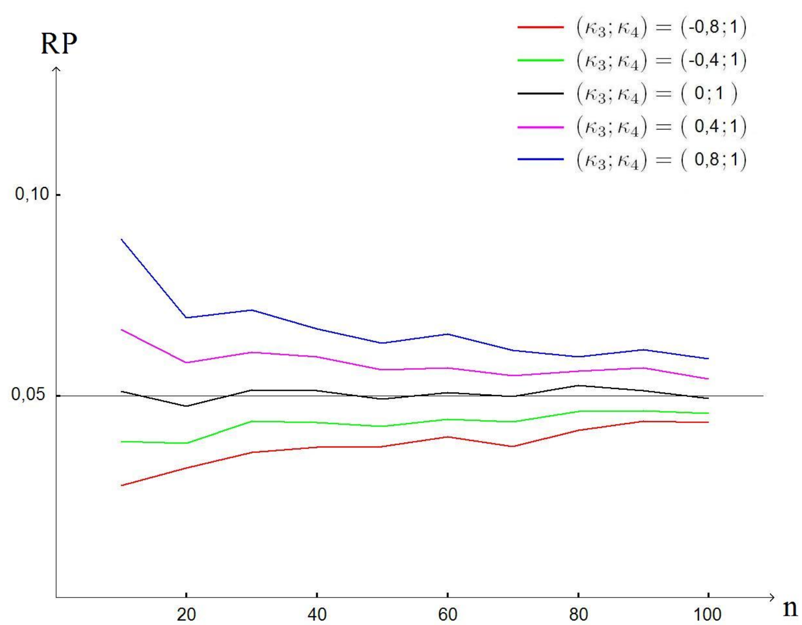

Figure 7.4: RP of parametric bootstrap tests with $\kappa_4 = 1$ and κ_3 varying.

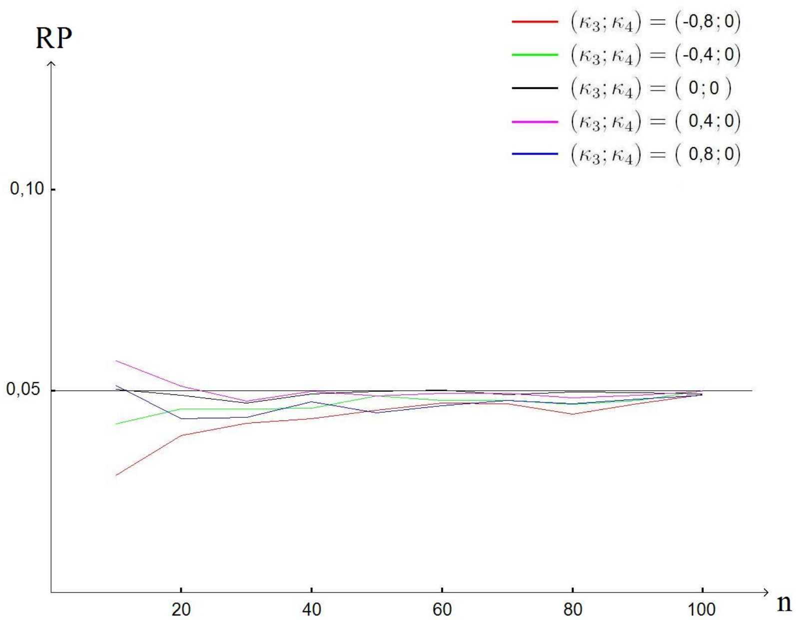

Figure 7.5: RP of non-parametric bootstrap tests with $\kappa_4 = 0$ and κ_3 varying.

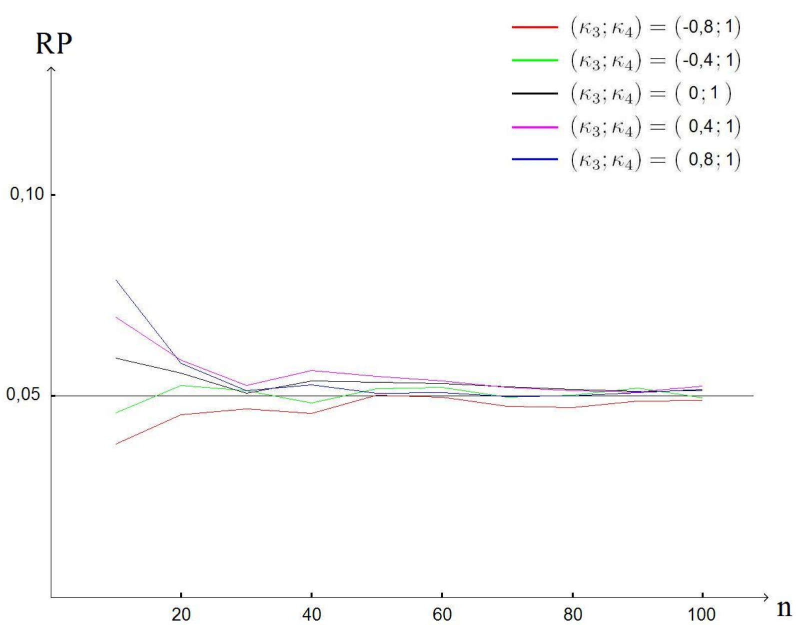

Figure 7.6: RP of non-parametric bootstrap tests with $\kappa_4 = 1$ and κ_3 varying.

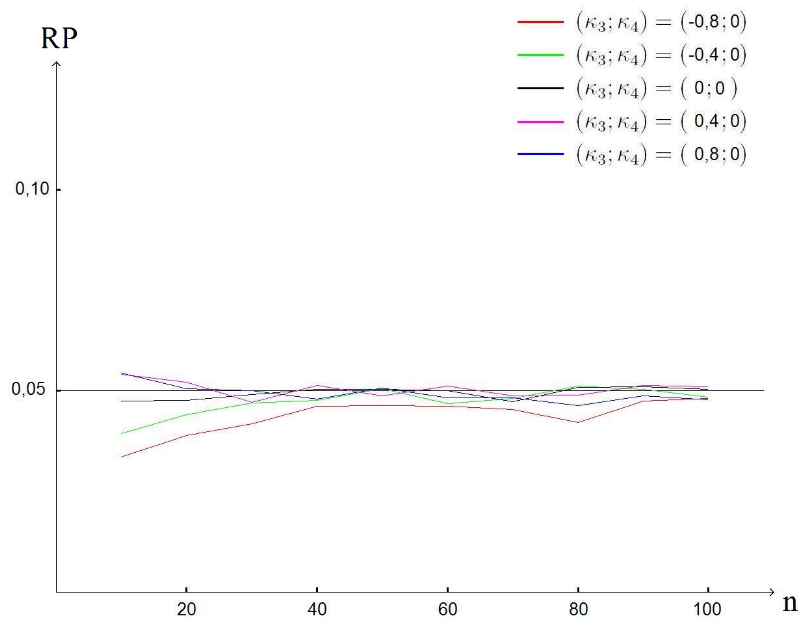

Figure 7.7: RP of parametric bootstrap CHM tests with $\kappa_4 = 0$ and κ_3 varying.

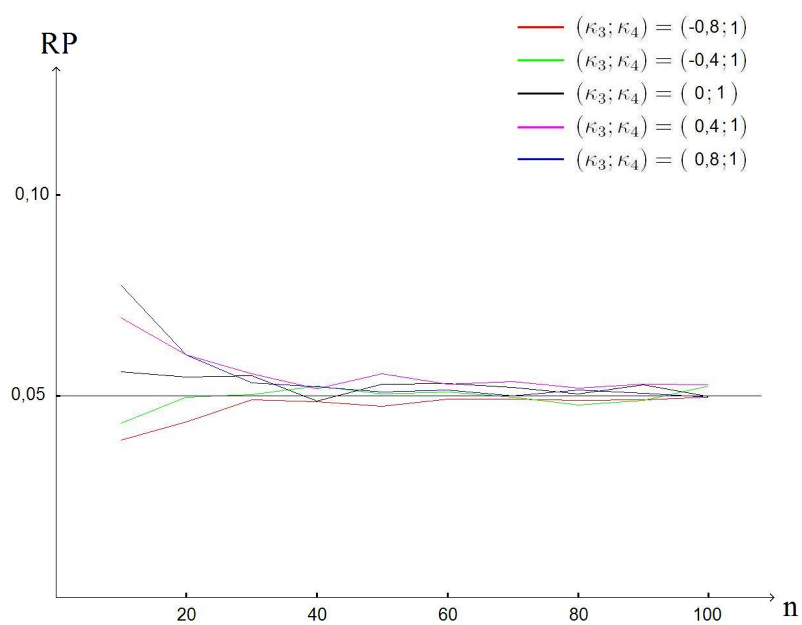

Figure 7.8: RP of parametric bootstrap CHM tests with $\kappa_4 = 1$ and κ_3 varying.

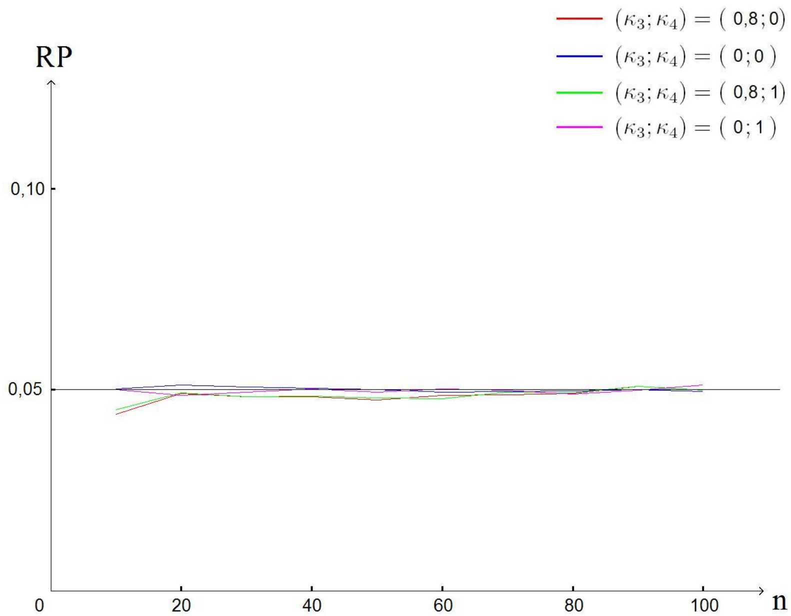

Figure 7.9: RP of the asymptotic tests.

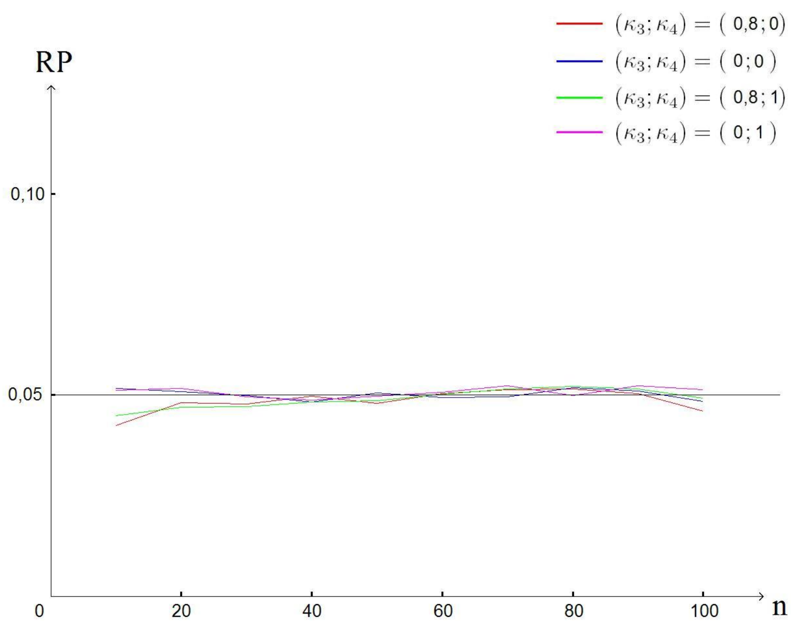

Figure 7.10: RP of the parametric bootstrap tests.

7.2 $y_t = \mu + \alpha_0 x_t + u_t$

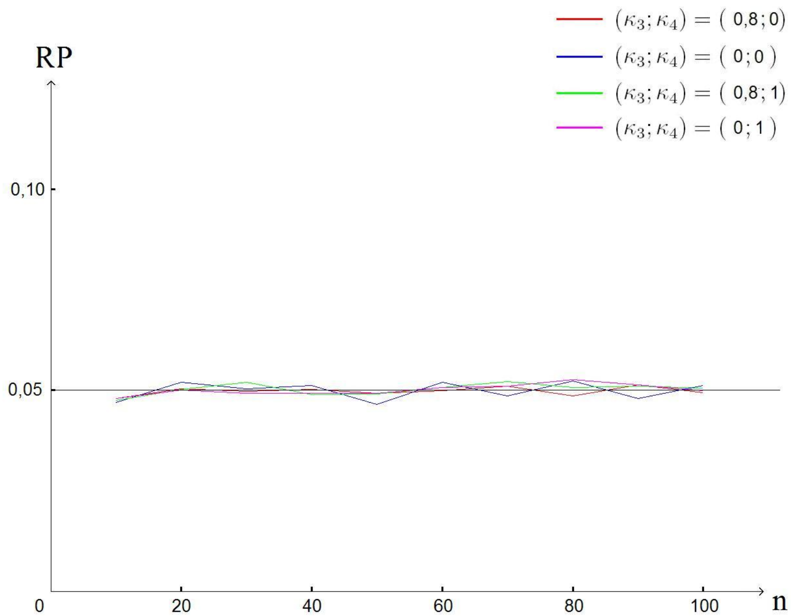

Figure 7.11: RP of the non-parametric bootstrap tests.

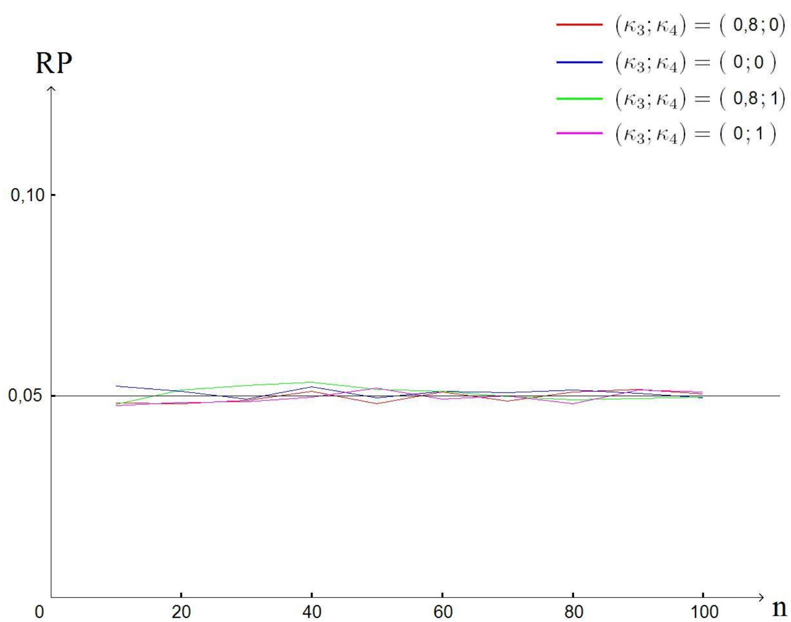

Figure 7.12: RP of the parametric bootstrap CHM tests.

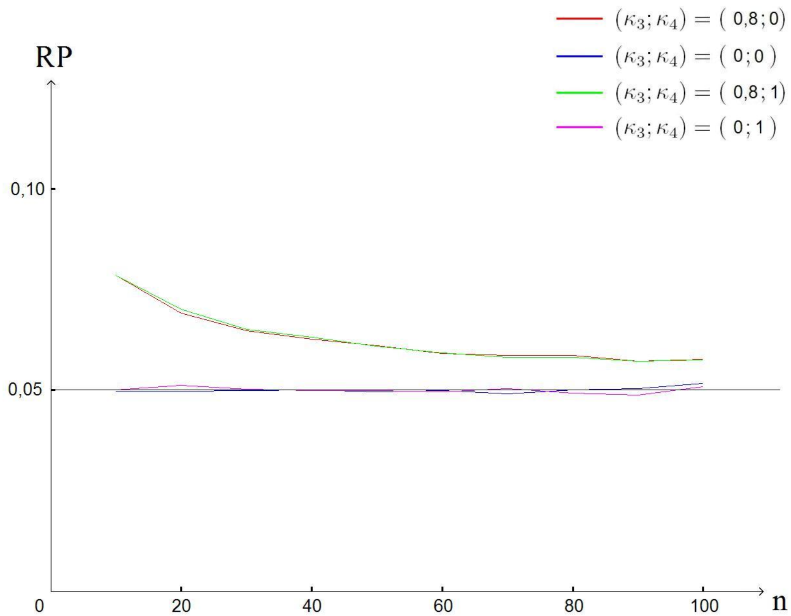

Figure 7.13: RP of the asymptotic tests.

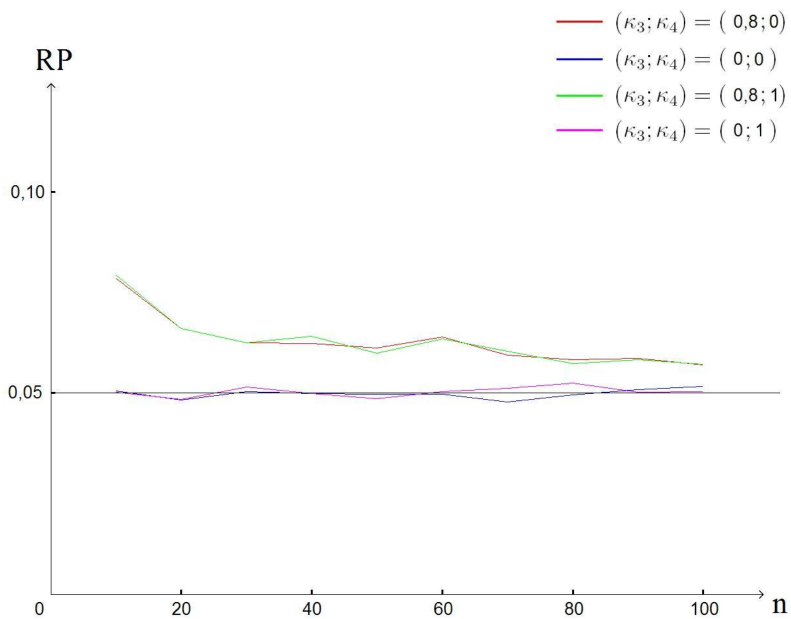

Figure 7.14: RP of the parametric bootstrap tests.

7.3 $y_t = \alpha_0 x_t + u_t$

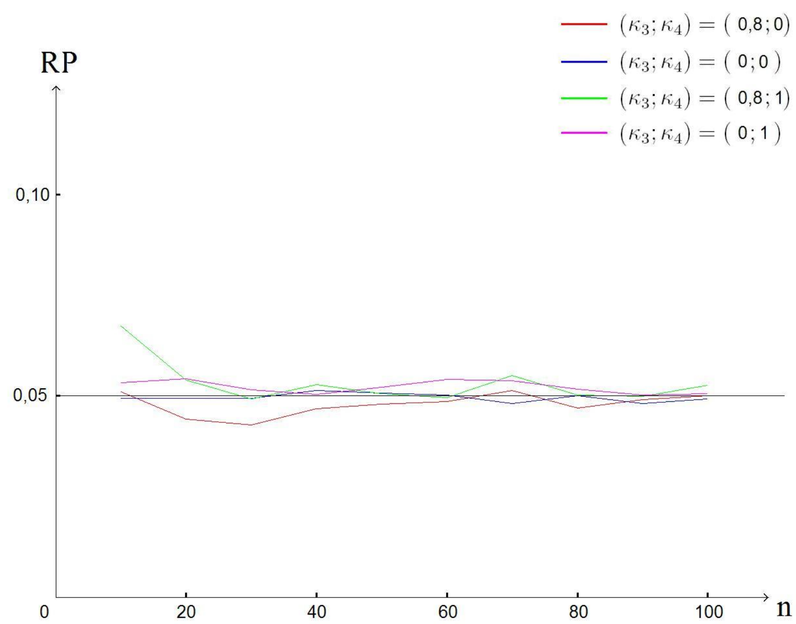

Figure 7.15: RP of the non-parametric bootstrap tests.

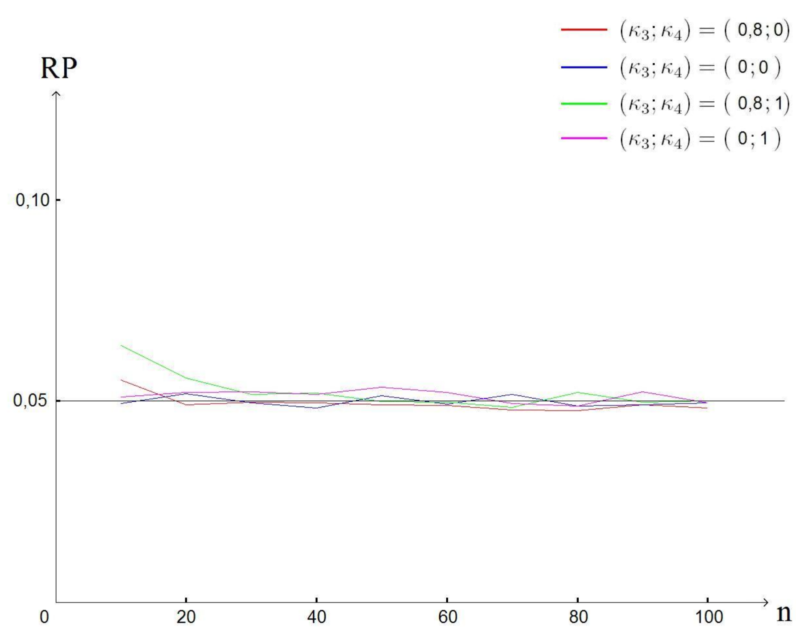

Figure 7.16: RP of the parametric bootstrap CHM tests.

7.4 Different variables.

We give all variables we use to compute the ERPs of this paper. Here, μ_i denotes the uncentered moment of the disturbances distribution at order i.

$$m_1 \equiv n^{-1}\sum_{t=1}^{n} x_t \quad \textit{with} \quad \lim_{n\to\infty} m_1 = O(1) \tag{7.1}$$

$$m_2 \equiv n^{-1}\sum_{t=1}^{n} x_t^2 \quad \textit{with} \quad \lim_{n\to\infty} m_2 = O(1) \tag{7.2}$$

$$m_3 \equiv n^{-1}\sum_{t=1}^{n} x_t^3 \quad \textit{with} \quad \lim_{n\to\infty} m_3 = O(1) \tag{7.3}$$

$$m_4 \equiv n^{-1}\sum_{t=1}^{n} x_t^4 \quad \textit{with} \quad \lim_{n\to\infty} m_4 = O(1) \tag{7.4}$$

$$w \equiv n^{-1/2}\sum_{t=1}^{n} u_t \quad \textit{with} \quad p\lim_{n\to\infty} w = N(0,1) \tag{7.5}$$

$$q \equiv n^{-1/2}\sum_{t=1}^{n} \left(u_t^2 - 1\right) \quad \textit{with} \quad p\lim_{n\to\infty} q = N(0, 2+\kappa_4) \tag{7.6}$$

$$s \equiv n^{-1/2}\sum_{t=1}^{n} \left(u_t^3 - \kappa_3\right) \quad \textit{with} \quad p\lim_{n\to\infty} s = N(0, \mu_6 - \kappa_3^2) \tag{7.7}$$

$$k \equiv n^{-1/2}\sum_{t=1}^{n} \left(u_t^4 - 3 - \kappa_4\right) \quad \textit{with} \quad p\lim_{n\to\infty} q = N\left(0, \mu_8 - (3+\kappa_4)^2)\right) \tag{7.8}$$

$$X \equiv n^{-1/2}\sum_{t=1}^{n} (u_t x_t) \quad \textit{with} \quad p\lim_{n\to\infty} X = N(0, m_2) \tag{7.9}$$

$$Q \equiv n^{-1/2}\sum_{t=1}^{n} \left((u_t^2 - 1)x_t\right) \quad \textit{with} \quad p\lim_{n\to\infty} Q = N(0, (2+\kappa_4)m_2) \tag{7.10}$$